Context-aware intelligent assistant approach to improving pilot's situational awareness

Lilly Spirkovska *

NASA Ames Research Center, MS 269-3, Moffett Field, CA 94035-1000

Suresh K. Lodha

University of California, Computer Science Department, Santa Cruz, CA 95064

Abstract

Faulty decision making due to inaccurate or incomplete awareness of the situation tends to be the prevailing cause of fatal general aviation accidents. Of these accidents, loss of weather situational awareness accounts for the largest number of fatalities. We describe a method for improving weather situational awareness through the support of a context-aware, domain and task knowledgeable, personalized and adaptive assistant. The assistant automatically monitors weather reports for the pilot's route of flight and warns her of detected anomalies. *When* and *how* warnings are issued is determined by phase of flight, the pilot's definition of acceptable weather conditions, and the pilot's preferences for automatic notification. In addition to automatic warnings, the pilot is able to verbally query for weather and airport information. By noting the requests she makes during the approach phase of flight, our system learns to provide the information without explicit requests on subsequent flights with similar conditions. We show that our weather assistant decreases the effort required to maintain situational awareness by more than 5.5 times when compared to the conventional method of in-flight weather briefings.

---

1 Introduction

Only 4% of non-military aircraft are classified as commercial air carriers. The other 96% are considered general aviation (GA). The target of our research is the 75% of civil aircraft that are characterized as small aircraft with 2 to 6 seats and 1 or 2 piston engines. Within this group, loss of weather awareness historically has the highest fatality rate (73% fatality rate in 2001 (Landsberg, 2001)). Two reasons have typically been cited as a cause or contributing factor: pre-flight weather briefings are in a format that is difficult to interpret and in-flight weather briefings are difficult to obtain and interpret. Much previous research has focused on the pre-flight briefing issue (National Weather Service, 2002; Ruokangas and Mengshoel, 2003; Scanlon, 1994; Uckun et al., 1999) and has resulted in graphical representations that are easier to interpret. Our current research focuses on the in-flight briefing issues.

The conventional method for in-flight briefings is for the pilot to obtain via aircraft radio a verbal update of conditions from a ground-based weather specialist. This process has a number of disadvantages. First, it is difficult to create a big picture of conditions from a verbal description. Next, a single specialist is responsible for communicating with a potentially large number of pilots. The number of pilots seeking information increases further when the weather is worst and updates are most needed. This can lead to long waits for access and to shortened interactions that allow the pilot to receive only very specific information. Last, and perhaps the biggest disadvantage, the pilot must initiate the request for weather updates. Unless

* Corresponding author.
Email addresses: spirkov@email.arc.nasa.gov (Lilly Spirkovska), lodha@cse.ucsc.edu (Suresh K. Lodha).

the weather in the local area is deteriorating, pilots often rely on the accuracy of the pre-flight briefing, dwindling their diversion options if they encounter unexpected conditions later in the flight.

Utilizing a human-centered methodology, we developed a weather briefing system to address the deficiencies of the conventional method. Because the in-flight environment is very vision intensive, any new technology must minimize required visual attention. Our system, Aviation Weather Environment (AWE), fulfills this requirement in four ways. First, it automatically checks for weather updates, freeing the pilot to manage essential flying tasks without distraction. Second, it provides three different graphical representations of important weather elements overlayed on top of a navigation map. The representations depict current and forecast conditions in an easy to interpret manner and are geographically positioned next to each applicable airport to enable the pilot to picture conditions along the route of flight. Third, it implements a speech-based user interface to enable the pilot to extract weather information verbally. Last, it implements an interface agent. The agent automatically checks for weather updates and warns the pilot of any conditions that do not meet her preset limits. It is context-aware, deducing the current phase of flight from the GPS location and the pilot-entered route of flight. It uses this context information to further filter the warnings that it issues. In addition, it tracks the pilot's verbal requests for information and learns her habits for the type of information of interest at particular times during a flight.

The representations are described in Spirkovska and Lodha (2002) and the speech-based user interface is described in Spirkovska and Lodha (2003). In this paper, we describe the weather agent and the learning algorithm. We also report on the methodology and results of an evaluation of whether AWE decreases a pilot's workload and whether pilots enjoy interacting with it. We conclude with possibilities for future enhancements.

2 Related Work

One of the more recent innovations in general aviation is the capability to receive up-to-date digital weather data while in flight (FAA, 2002; Horne, 2002; McClellan, 2003; Ruley, 2002). Two main approaches have emerged for this so-called data link capability (AirCell, 2002; ARNAV, 2002; EchoFlight, 2002; Honeywell, 2002). One utilizes ground station networks and continuously broadcasts weather products; the other relies on low-Earth-orbit communication satellites to respond to explicit pilot requests for individual weather reports. The method of pilot interaction with each system differs due to the approach of transmission (Collins, 2003). In the broadcast approach, the pilot is provided with all weather reports for a large geographic area. Most products are broadcast once every five minutes. In contrast, the request-and-reply approach requires that the pilot request an individual report (e.g., current conditions at SFO airport). After a delay (which can vary from one to twenty minutes), the report is transmitted to the pilot.

In both cases, the pilot has access to NEXRAD radar reports, current weather reports, and forecast weather reports. NEXRAD radar reports are shown graphically, forecast conditions are shown textually, and current conditions can be shown either textually or graphically. The graphical depiction of current conditions shows a very small portion of the available data; to get the full report, the pilot must read the textual report. In any case, the weather reports are provided to the pilot to filter for appropriateness to route-of-flight and to interpret without any automated assistance.

In contrast to the above methods, research is underway on automated assistance to improve pilots' situational awareness (Ballin et al., 2002; Endsley, 1997; Olson and Sarter, 1998; Painter et al., 1997; Ruokangas and Mengshoel, 2003; Uckun et al., 1999). One such decision support system, AWARE, aims to help pilots interpret weather reports and is mostly closely related to ours (Ruokangas and Mengshoel, 2003). In-flight, AWARE monitors the proximity of hazards such as thunderstorms and informs the pilot visually by color-coding the corresponding button on a graphical user interface. The pilot can then drill-down to textual information or examine NEXRAD radar data to fully understand the situation. Our system differs in the weather elements we consider, the additional flexibility and support we offer the pilot to filter the available data, and the multimodal approach we use to inform her of relevant anomalies.

3 Weather Agent

AWE's weather agent is a context-aware, task-knowledgeable, personalized, and adaptive assistant. The task of the weather agent is to provide relevant in-flight weather data to the pilot at a relevant time with or without an explicit pilot request. The weather agent has built-in knowledge of the domain, the task, and the pilot. Domain knowledge is used to determine if a parameter or a trend is cause for concern. Task knowledge is used to determine if that data is relevant in a particular phase of flight. And pilot knowledge is used to determine when and how to provide relevant data to the pilot without being obtrusive and bothersome. Based on an analysis of the pilot's workload and need for weather updates during different phases of flight, we limit the scope of the agent's capabilities to the cruise and approach phases. All of AWE's capabilities are available via direct manipulation during all phases of flight, but unsolicited suggestions are offered only during cruise and approach.

3.1 Domain Knowledge

Wind, visibility, cloud height, temperature, turbulence, and icing can each adversely affect a flight. Wind, visibility, and cloud height are most often considered a cause or factor in general aviation aircraft accidents (Keel et al., 2000). Airport-specific current condition reports, airport-specific forecasts, and winds aloft reports are most frequently consulted by pilots (Keel et al., 2000) to obtain data about winds aloft, surface winds, cloud height, visibility, temperature, and dew point. AWE collects these three reports about all route airports, from departure through destination. It then filters the data using built-in knowledge of which elements affect pilot's decisions. All other data can safely be ignored, though it is available for the pilot to read if desired. In addition to weather data, AWE uses data about the airport layout (e.g., orientation of runways and the traffic pattern) to further assist with decisions regarding landing.

3.2 Task Knowledge

Although domain-based filtering decreases the amount of data significantly, not all of the remaining data is relevant to the task. AWE uses task knowledge coded in a rule-base to further reduce unnecessary data.

During cruise, the pilot may be interested in (Keel et al., 2000): the local altimeter setting; changes in winds aloft if the destination is no longer within fuel range constraints; significant discrepancies between the current conditions and the conditions forecast for the time period; current conditions at the destination and en-route airports that do not meet her constraints; and revised forecasts that do not meet her constraints. Each constraint can be specified by the pilot, as described below. Some constraints, such as visibility, can be checked against incoming data directly. Other constraints require multiple sources of data. For example, checking crosswind constraints requires incoming data and knowledge of the airport layout. Similarly, correlation between actual and forecast conditions requires inspecting multiple incoming data elements. Checking these compound constraints requires the most pilot cognitive effort. AWE uses rules and heuristics to assist the pilot by automating the checks.

During the approach phase, determined from GPS location and the flight plan, the pilot may be interested in destination area conditions, including visibility, ceiling, surface wind, and density altitude; the airport layout including runway orientation, field elevation, pattern altitude, location of the pattern (i.e., left or right traffic pattern), and runway length; and altimeter setting for the destination area. If conditions are deteriorating, she may want to know the nearest airport with acceptable conditions. Even when conditions at the destination airport are acceptable, the proximity of adverse conditions may also influence her expectations for landing there. As described below, AWE attempts to assist the pilot by tracking the above information and informing her only when appropriate as determined by encoded pilot preferences and relevance to the current phase of flight.

3.3 Pilot Knowledge

Each pilot's go/no-go decision is influenced by her own characterization of the weather conditions. For example, one pilot's comfort level is exceeded with a 15 knot surface wind; another pilot is not concerned until the wind exceeds 30 knots. Additionally, each pilot's desire for automated alerts differs. Whereas one pilot may want automatic alerts if the visibility at her destination decreases below 10 miles, another may prefer to query manually until it decreases below 4 miles, and yet another may be annoyed by interruptions and would rather always query manually. The desired form of alerts also varies. Some pilots dislike verbal messages - they are content to fly along listening to air traffic controllers (ATC), other pilots, or their passengers and not have a computer talking at them. Others find that they prefer to look out the window and may miss information unless it is provided aurally. AWE provides personalized assistance by knowing each pilot's definition of acceptable weather conditions and *if*, *when*, and *how* to volunteer information about deteriorating conditions.

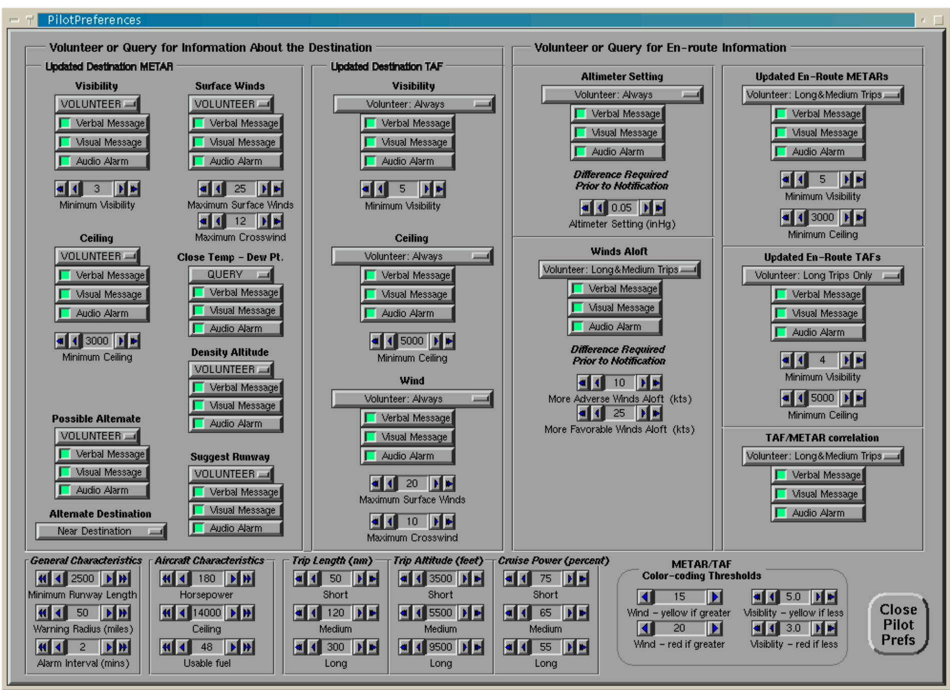

Fig. 1. Pilot preferences window. The pilot can modify stereotypical values to better reflect her preferences.

Defining Acceptable Conditions Weather influences each pilot's continue/land-now en-route decision differently. To accommodate different experience or comfort levels, each pilot can personalize AWE to warn her of conditions she considers adverse. Via the graphical user interface shown in Figure 1, the pilot can set thresholds for any of a large set of situations. For example, she can specify that she wants to be informed of any revised forecasts that predict crosswind values at the destination airport greater than 15 kts. In addition to weather elements, the pilot can specify minimum acceptable runway length; whether AWE should select alternate airports near the current location or near the original destination; performance characteristics (horsepower, maximum ceiling, and usable fuel) of the aircraft; typical cruising altitude and power settings for short, medium, and long trips; and a definition of short, medium, and long trips. Because of the multitude of options and to obviate the need for each user to personalize it prior to use, AWE begins with limits reasonable for a typical pilot, as suggested by Elaine Rich's (Rich, 1998) stereotypical user model. The pilot can then adjust the limits as desired.

Defining Notification Conditions For an individual pilot, the type of assistance desired may vary with the length of a flight. For example, if a pilot always departs with at least enough fuel to easily complete a medium length trip, she may be concerned about winds aloft changes only on a long flight. On the other hand, she may always want the most recent altimeter setting, even on local flights. To accommodate individual preferences for automatic notification, AWE maintains *when* the pilot wants to be informed about each condition defined by the eight combinations of the cross product of *(short, medium, long)* crossed with *(volunteer, query)*. *Volunteer* and *query* refer to whether AWE volunteers that it has detected conditions outside the stated limits, or whether it waits for the pilot to query for the information. Similar to the limits, preferences for notification are initialized for a stereotypical pilot and can be adjusted to suit an individual pilot.

Defining Notification Methods In addition to *when* the pilot wants to be informed, AWE maintains a list of *how* she wants to be informed about each condition: verbally, visually, or with an audio alarm. AWE is initialized to use all three options for each condition, but a pilot can adjust this to suit her preferences. Both visual and verbal warnings are terse: they provide only enough information to make the pilot aware of the conditions; additional details can be extracted via direct manipulation. Because the pilot may be busy with other tasks when warnings are issued, a historical record of visual warnings is available in a separate *AWE messages* scrolling window, shown in Figure 2.

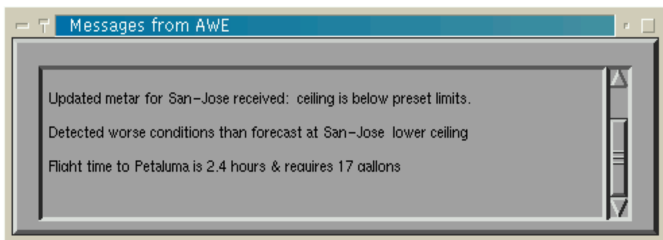

Fig. 2. AWE Messages window. The pilot can scroll through a history of messages provided by AWE.

After extended interaction with a particular pilot, AWE may determine that the above notification items are inadequate for her. It may detect that she often asks for additional information or asks for volunteered information at additional points along the flight, not just when new weather reports are received. In this case, AWE expands the set of information it provides automatically by applying machine learning techniques to learn the pilot's habits, as described in the next section.

4 Automatic Adaptation

Many pilots develop habits for when they want certain information. For example, a pilot may have a habit of looking up information about the airport when within 30 miles, getting the destination surface wind and verifying the runway orientation when within 20 miles, and then verifying the pattern altitude when within 10 miles. When two pilots fly together often, even though the second pilot is not a required crew member, he is often used as an additional resource, especially when the pilot's workload increases during the approach phase. After enough flights together, the second pilot may learn to anticipate when the pilot will ask for his assistance.

AWE implements similar anticipation by tracking what type of information the pilot asks for (e.g. pattern altitude, nearest IFR conditions, etc.), and when (distance from the destination and weather conditions at the destination) she asks for it. AWE fulfills the request and adds a vector representing the directive to its set of directives used to learn the pilot's habits. After five requests for a particular type of information, AWE generalizes the habit using a modified reflex learning technique described below. When similar conditions are encountered, AWE volunteers the information without prompting. As an example, if she asks for nearest IFR conditions when she is within 50 miles from the airport and the temperature-dew point spread is 3 degrees Celsius, AWE may learn that it should volunteer that information without prompting. However, if the training set for *IFR conditions* consists of (density altitude, temperature-dew point spreads) of $\{\ldots$, (2500', 3), (500', 3), (1000', 2), $\ldots\}$, ignoring all the other elements, it may also learn that if the density altitude is less than 2500', it should volunteer nearest IFR conditions. Like a human assistant, AWE improves its understanding of the pilot's habits with additional training. Also like a pilot assistant, it uses domain knowledge to eliminate irrelevant generalizations.

Learning Algorithm Details There are various methods for teaching machines how to extract commonalities between situations in order to make decisions using resulting generalizations. Typically, a set of example situations is encoded by a set of properties and the learning algorithm must formulate general descriptions for related subsets of situations. When a new situation is presented to the system, it uses these descriptions to determine what type of situation most closely resembles it. Using this classification, it could then decide what to do next.

Each of AWE's example situations is encoded by the ten-tuple (type of information requested, distance to destination, trip length, visibility, ceiling, wind, crosswind, temperature, temperature-dew point spread, density altitude) where the last seven elements represent the weather conditions at the destination airport. By using the *type of information requested* property as the classification for the example, we could apply a decision tree (Russell and Norvig, 1995) or a neural network (Rumelhart et al., 1986) algorithm to extract the type of conditions that prompt the pilot to seek that type of information. Or we could apply a Bayesian classification system (Cheeseman and Stutz, 1996) directly on the 10-element vectors and have it separate the examples into related subsets that describe the pilot's habits. However, all three techniques require a large set of training examples to enable relevant learning. Without adequate training data, the system could easily extract the unlikely rule mentioned previously: density altitude of less than 2500' prompts the pilot to request *IFR conditions*.

To make AWE's learning algorithm effective with less training data, we chose to integrate domain specific knowledge with a reflex, or stimuli-response, learning approach (Russell and Norvig, 1995). Although the pilot may ask for various items during the approach and may ask for them to be provided using the three different formats, the AWE prototype only tracks the pilot's habits with eight major approach-relevant directives and only when she asks about the destination airport. Each time one of these directives is issued, the 10-property description is added to the training set, and AWE is retrained to include the new example in determining how far out and under what weather conditions AWE should volunteer a response without prompting. To reduce the likelihood of misinterpreting a non-habitual directive as a habit, generalization occurs for a *type of information* only if there are at least five examples.

The learning algorithm, Enhanced Algorithm for Reflex Learning (EARL), begins by separating training examples by type of information requested. Each subset is learned separately. Reflex learning relies on memorizing each training example to determine future action under the same conditions. EARL enhances this approach by generalizing from the given examples to determine future action under similar but unseen conditions. It accomplishes this by fusing the values for each property using a union, minimum, maximum, or averaging operation depending on the property. *Trip length* is fused by taking the union of the values evident in the examples. Thus, if the examples represent directives issued on three short and four medium length trips, the fused value will be (small or medium), but not long. *Distance from destination* is fused by taking an average of the example values. *Visibility*, *ceiling*, and *temperature-dew point spread* are fused by taking the maximum value for each property, and the remaining properties (*wind speed*, *crosswind*, *temperature*, and *density altitude*) are fused by taking their respective minimum value. Hence, EARL may learn to provide a particular type of information when the visibility is less than 20, wind is greater than 15, or the temperature is greater than 80 degrees Fahrenheit. EARL further enhances the reflex learning approach by applying domain knowledge: EARL is provided with a list of relevant properties for each *type of information* requested. For example, for *say density altitude* requests, only the *temperature* and *density altitude* properties likely contribute to a pilot's decision to make the request. This prevents it from learning from irrelevant coincidences of data combinations. Finally, learning based on a small set of examples could lead to the undesired extraction of unforeseen habits. AWE provides a graphical method, shown in Figure 3, to enable the pilot to view, modify, and repress learned rules to better reflect her desired interaction.

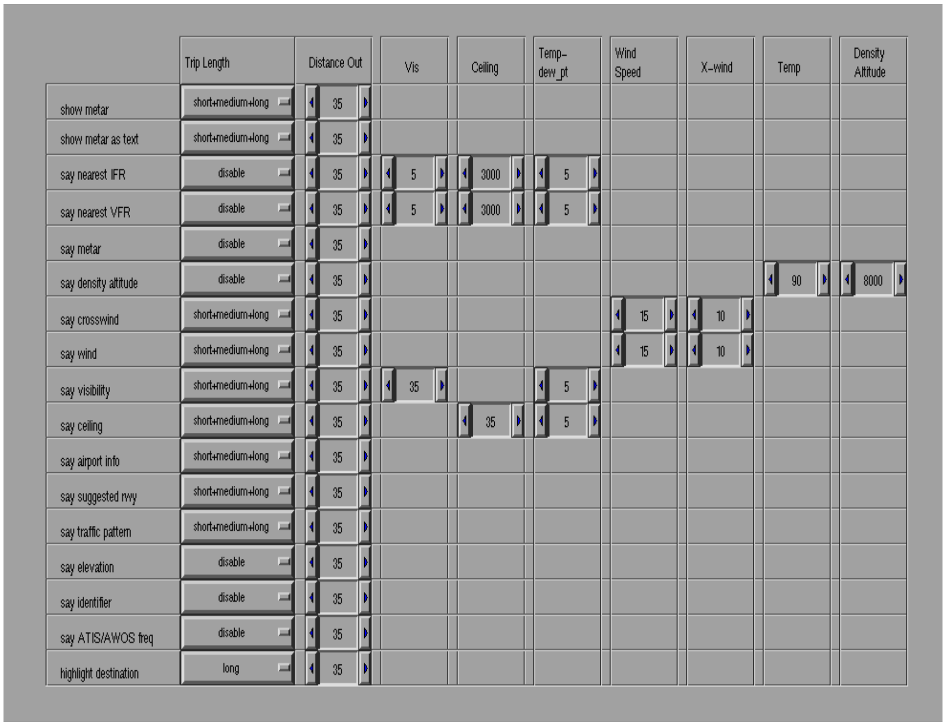

Fig. 3. AWE Learned Habit Modification Window. The pilot can view, modify, or repress habits learned by AWE to better reflect her information needs.

5 Workload Reduction Evaluation

The design of AWE was influenced by the first author's experiences as a general aviation pilot. In addition, feedback was taken on many issues at several stages from different pilots to ensure that the system remains pilot-friendly and usable. Finally, the system was more formally evaluated by pilots. The evaluations focused on determining whether pilots found AWE to be a useful tool and whether it decreased their workload. The representation aspect of AWE was compared with several other weather data visualization systems (National Weather Service, 2002; Pruyn and Greenberg, 1993; Scanlon, 1994; Uckun et al., 1999; The-Weather-Channel, 2002) and the pilots' opinions were elicited via a questionnaire and short interview. The results of the representation evaluation are presented in Spirkovska and Lodha (2002). The weather agent aspect was evaluated using a part-task simulation and a post-simulation questionnaire to obtain quantitative workload reduction values and pilots' subjective opinions. We describe the process and results of the workload evaluations below.

The primary question of the agent evaluation was *Does AWE reduce the pilot's weather-related decision making workload?* Specifically, we wanted to answer *Is the time to make a decision decreased by AWE?* Related follow-up questions were *If it does, how? If not, how can it be improved?* Although the weather agent is designed primarily to provide assistance in flight, we include a brief description and results of the evaluation of AWE for pre-flight tasks for completeness.

Pilots Six general aviation pilots participated in the workload reduction studies. The approximate total flight time of the pilots ranged from 160 hours to 3200 hours with a mean of 1326 hours and a median of 650 hours.

Pre-flight To measure preflight workload reduction, we asked each pilot to interpret the weather for a pre-defined route. The pilots were given pertinent information about the aircraft. Additionally, to eliminate variations due to different pilots' definition of acceptable weather, they were instructed to assume a set of pilot limits for visibility, ceiling, and surface wind speed. Prior to beginning the study, we provided the go/no-go decision questions we would ask at the conclusion of the study. Finally, we answered any logistic questions they had.

Each pilot was provided with two routes, counterbalanced in the type of weather conditions, number of en-route report elements, and distances. Two sets of airports were selected to decrease possible effects of transferring learned knowledge about the airports and regions between tasks. Each pilot was assigned two tasks. For both tasks, we measured the amount of time required to answer the go/no-go questions. For the first task, the pilots were provided a standard computer weather briefing for a 233 nm flight. For the second task, the same pilots were provided with an AWE briefing for a different 223 nm flight. The AWE briefing consisted of the pilot selecting desired values for departure time, cruising altitude and airspeed and then viewing current and forecast conditions along a pilot-specified route or the area of interest. An example view of an AWE display of current conditions and winds aloft for a four-airport route is shown in Figure 4. It shows two of the three possible representations - textual and symbolic. An example view utilizing the triangular icon representation to display forecast conditions for a particular region is shown in Figure 5. More details on how to interpret the figures are available in Spirkovska and Lodha (2002).

Answering the questions using the conventional briefing required an average of 60.2 minutes, whereas answering them using AWE required an average of 23.5 minutes. Hence, answering preflight relevant questions is more than 2.5 times faster with AWE. Additionally, the AWE times include the time to get comfortable with the system.

In-flight For the in-flight interaction evaluation, due to constraints with aircraft, data transmission, and pilot resources, we measured workload reduction in a ground-based part-task simulation. The same six participants were randomly divided into two groups. Half of the participants had access to the conventional method of in-flight weather updates - known as *Flight Watch* - and the other half had access to AWE updates. The participants were all presented with a pre-assigned route of flight and a pre-flight briefing. The weather conditions were forecast to be ideal for

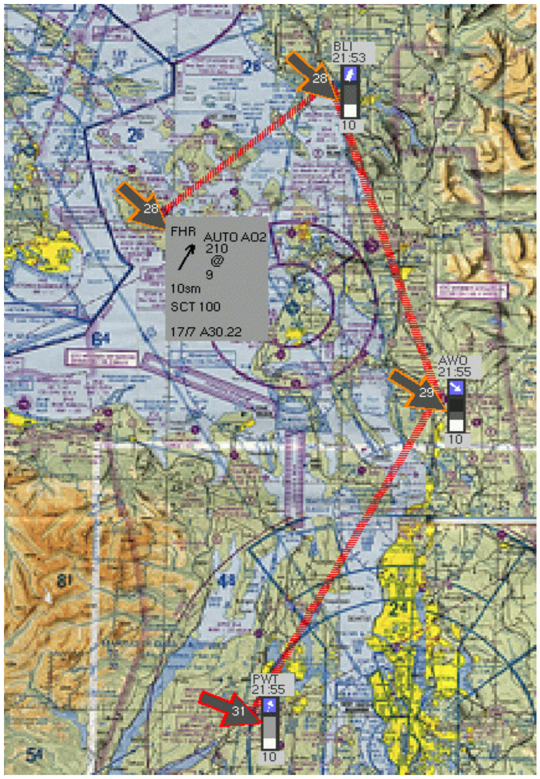

Fig. 4. Route-specific current conditions and winds aloft shown alongside a pilot-selected route. Wind velocity at the pilot selected altitude is depicted graphically with a black arrow. The airport conditions are shown using the symbolic representation (the vertical rectangle) and the textual representation. The symbolic representation shows, top to bottom, surface wind velocity, cloud coverage at altitudes from 12,000 feet down to the surface, and surface visibility. The textual representation shows all the elements available in computer briefing reports except for remarks.

the first half of the flight and marginal but acceptable for the second half. The assigned route required four hours of flight time but we started the simulation already established in cruise 1.5 hours into the flight. The pretense of flying the first half allowed us to simulate restricted access to weather updates - rather than beginning the simulation with the most recent weather reports, they began with two hour old reports. This increased the likelihood of getting an in-flight weather update. To ensure an update request, we forced a diversion: they were informed that the clouds

Fig. 5. Area-wide forecast conditions display using triangular icons. The top, lower left, lower right, and the middle subtriangles represent winds, visibility, clouds and temperature/dew point spread conditions respectively. Red, yellow, white, and gray colors indicate alert, caution, normal conditions, and unavailable respectively.

were too low to continue using the original route of flight. However, to evaluate the effectiveness of AWE during the approach phase, after they planned for the diversion, they were informed that the clouds cleared along the original route and to continue to the original destination. We discuss the measurements taken in a later subsection.

To further decrease the demand on the participants' time, we accelerated simulated time by a factor of six, thereby decreasing the simulated flight time from 2.5 hours to 25 minutes. The timing of issuance of weather reports was accelerated to maintain consistency.

To simulate the workload of a flight, we designed a simulated cockpit, shown in Figure 6. In the real world, pilots must fly the aircraft, maintaining desired altitude, airspeed, and heading; scan the engine instruments, looking for anomalies; scan for traffic, maneuvering to maintain separation; and maintain weather awareness. The virtual cockpit simulates these to varying degrees. The participants were provided with complete instructions on how to fly the route and deal with traffic and engine anomalies.

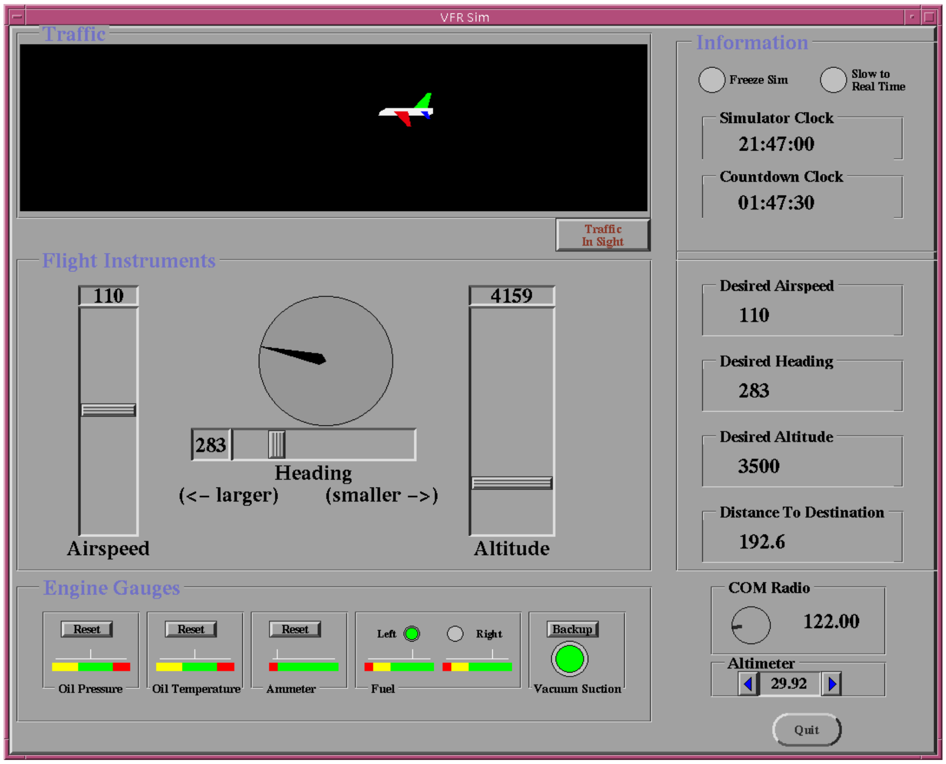

Fig. 6. AWE Evaluation Simulator. Used to simulate workload of a VFR flight. The pilot must look for traffic; maintain desired airspeed, heading, and altitude; monitor engine parameters; manage fuel; and tune to appropriate weather frequencies for weather and altimeter setting updates.

Measurements With both sets of pilots busy "flying" the aircraft and scanning for traffic and engine anomalies, we compared their level of weather awareness as well as the workload required to maintain it. To eliminate having them work toward a task, the participants were not aware of the parameters we were measuring. Specifically, we measured the amount of time spent getting updated weather conditions, how quickly they formulated an alternate plan for the forced diversion, and how often they updated the altimeter setting along with the time required to do so. We also measured their performance in detecting the anomalies presented by the virtual aircraft, how quickly they noticed traffic, and how quickly they returned to the desired altitude, heading, and airspeed. These quantities provided a measure of whether they ignored the flying task in preference to gathering weather data. In addition, the pilots getting AWE updates were asked about their satisfaction with AWE and suggestions for increasing its usefulness. The pilots using Flight Watch were given a demonstration of AWE and asked the same questions.

Additional Logistics Because we were doing a ground-based simulation, using simulated weather data, and an accelerated clock, Flight Watch was simulated. The role of Flight Watch was filled by another pilot who could be contacted via telephone (to simulate some of the non face-to-face communication issues). We did not simulate possible delays due to Flight Watch filling other pilots' requests. Hence, the measurements for the pseudo Flight Watch group represent a lower bound. The pilots were not restricted in the type of questions they could ask nor in the level of processing and analysis they wanted Flight Watch to provide. The only requirement was that they continue to fly the simulated aircraft while speaking to pseudo Flight Watch.

Communication with AWE was also assisted. The recognition accuracy of the speech recognition system we used (IBM ViaVoice) increases with speaker training - that is, speaker-dependent recognition accuracy is higher than speaker-independent accuracy. The grammar, though designed to mimic other pilot communication, requires some learning effort by the pilot. Testing for learnability requires repetitive usage. Instead, we chose to use available time with pilots to measure other aspects. To eliminate extra pilot effort required to train ViaVoice and to learn the grammar, we assisted the pilot by being a conduit for verbal requests to AWE. The simulation assistant (who previously trained ViaVoice) issued the spoken directives to AWE.

The pilots had two monitors directly in front of them. One monitor displayed the simulated cockpit while the other provided access to AWE. The pilots could use direct manipulation techniques (using either the GUI or speech user interface) to obtain current and forecast conditions and also rely on the weather agent to notify them (verbally, visually, or with an audio alarm) of updated conditions that may affect their flights. Although the route of flight was pre-selected, the pilots were not restricted in which airports they could seek information about or the type of information desired.

Each evaluation session lasted approximately 60 minutes and included time for familiarization with the virtual cockpit, for a preflight weather briefing, for a brief familiarization with AWE, for flying the flight, and for debrief. The AWE familiarization took the form of a demonstration for the Flight Watch group and was presented at the end of the session.

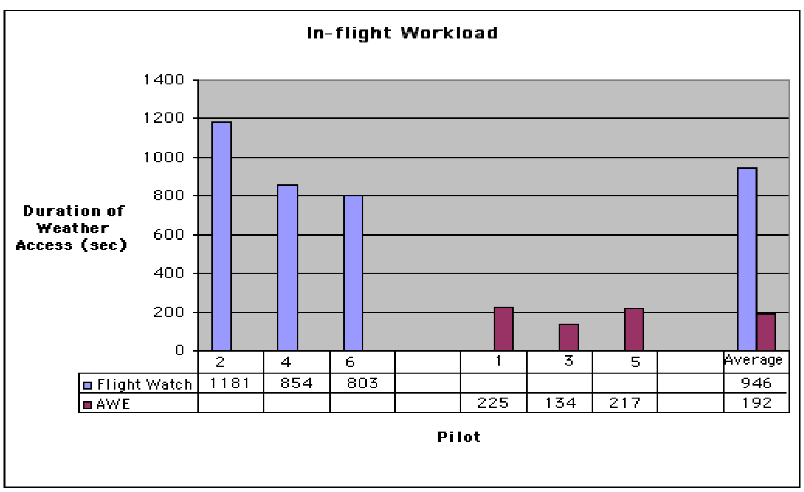

Fig. 7. Results of in-flight workload reduction user study. All times shown are in seconds.

Results The results of the in-flight evaluation are shown in Figure 7 and Figure 8. On average, the pseudo Flight Watch group spent 946 seconds tracking the weather,

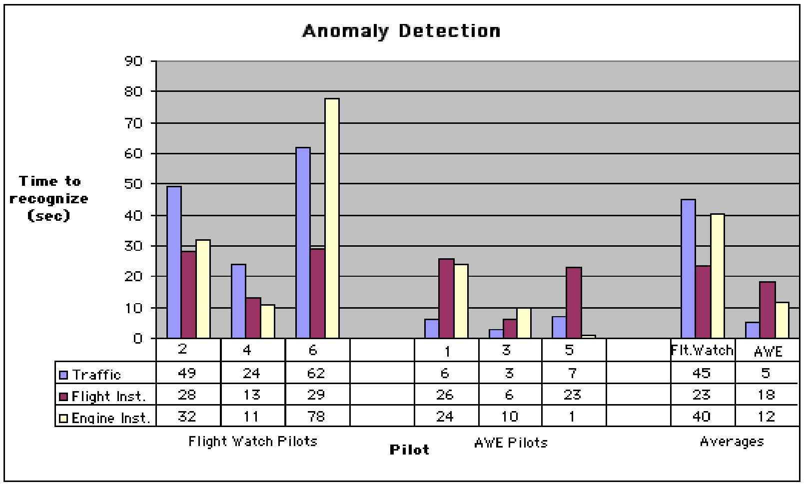

Fig. 8. Results of anomaly detection during in-flight workload study. Recognition times, in seconds, are shown for the three major anomaly categories. Pilots 1, 3, and 5 used AWE for weather updates, whereas pilots 2, 4, and 6 used Flight Watch.

mostly to select an alternate; detected the three major categories of anomalies (traffic, flight instruments, and engine instruments) in 36 seconds; and spent a couple of minutes formulating a landing plan. In general, they updated their altimeter setting after approximately 30 simulated minutes of flight and once again prior to landing.

On average, the AWE group spent 192 seconds tracking the weather, again mostly to select an alternate; detected the three major categories of anomalies in 12 seconds; and spent less than one minute formulating a landing plan. They each maintained a current altimeter setting within the precision defined by the default pilot preferences (0.02).

In the post simulation questionnaire, the AWE group was asked about the interaction experience. Specifically, they were asked how much they liked or disliked the automatic alerts about deteriorating weather, the completeness of the set of preferences they can specify and suggestions for others, how much they liked that it could learn their habits, and suggestions for improvements. The pseudo Flight

Watch group was given a demo of AWE after their simulations and asked their opinion on the same questions. The number of interactions for each participant was not adequate to enable AWE to learn his habits. Thus, the effect of the learning was experienced using the habits of the simulation assistant.

On average, the pilots gave AWE a score of 1.5 (where 1 is *liked it a lot* and 5 is *did not like it at all*) for how much they liked automatic alerts, 1 for preference completeness, and 1.8 for learning their habits. For automatic alerts, the two pilots who gave it a score worse than 1 (score of 2 and 3, respectively) cited that they do not necessarily want that much information, especially for en-route airports. We pointed out that they could set the pilot preferences to not volunteer en-route information or to conditions they would want alerts for rather than our very conservative default values. Habit learning was rated worse than 1 by three pilots. One was concerned about privacy issues - he did not want his personal habits for airport information retrieval stored and possibly revealed. Another indicated that he would like such information at unfamiliar airports, but not at the airports he flies from regularly since he already has that information memorized. The third pilot did not specify a reason. Pilot preferences were considered complete by nearly all the pilots. The only one who rated it a 2 did not offer suggestions on missing items. We expect these numbers and opinions to change with extended use of AWE and especially with in-flight use where the incentive to stay aware of weather changes is significantly increased. With extended use, they would no doubt find the features they like or dislike and the features that need to be added.

6 Future Enhancements

After a single use, our pilot evaluators suggested improvements in habit tracking, displays, overlays, and additional functionality. The evaluations studies demonstrate the workload reduction possible with use of the weather agent. They also resulted in suggestions for further improvements, including automatically extrapolating the values of preferences based on conditions observed during a series of flights; notifying the pilot which preferences she habitually violates; overlaying other types of restrictions (e.g., active military training routes, active military operation areas, temporary flight restriction areas) to help with route or diversion planning; and automatically highlighting suggested alternate airports and providing initial heading from current position.

In addition to these features, additional research can focus on utilizing more capable learning algorithms to expand what information can be learned. Moreover, we could further improve pilot's situational awareness by integrating AWE with terrain information, traffic information, navigation information, and aircraft health information. Individual agents responsible for each system separately could collaborate to determine the best course of action. Finally, the presentation of information and advice can utilize additional methods, such as a head-mounted display, tactile feedback, or data sonification.

References

AirCell, 2002. Airborne NEXRAD weather, www.aircell.com.

ARNAV, 2002. www.arnav.com.

Ballin, M. G., Sharma, V., Vivona, R. A., Johnson, E. J., Ramiscal, E., August

2002. A flight deck decision support tool for autonomous airborne operations. In: Proc. of Guidance, Navigation and Control Conference, Monterey, CA. Vol. AIAA-2002-4554.

Cheeseman, P., Stutz, J., 1996. Bayesian classification (AutoClass): Theory and results. In: Advances in Knowledge Discovery and Data Mining. AAAI Press/MIT Press.

Collins, R. L., April 2003. Weather on the panel: Working with the information after it arrives. Flying , 74–79.

EchoFlight, 2002. Data link weather, www.echoflight.com/Data_Link/data_link.html.

Endsley, M., 1997. Supporting situation awareness in aviation systems. In: IEEE International Conference on Systems, Man, and Cybernetics. Vol. 5. pp. 4177–4181.

FAA, 2002. Federal aviation administration flight information services, www.faa.gov/aua/ipt_prod/FISDL.

Honeywell, 2002. www.bendixking.com:80/static/fis/datalinkweather.html.

Horne, T. A., March 2002. Weather links. AOPA Pilot , 101–104.

Keel, B. M., Stancil, C. E., Eckert, C. A., Brown, S. M., Gimmestad, G. G., Richards, M. A., June 2000. Aviation weather information requirements study. Tech. Rep. NASA/CR-2000-210288, NASA.

Landsberg, B., 2001. Nall report: General aviation accident trends and factors for 1999. Tech. rep., AOPA Air Safety Foundation, www.aopa.org/asf.

McClellan, J. M., April 2003. Datalink weather: How Bendix/King and Garmin make every pilot's wish come true. Flying , 72–79.

National Weather Service, 2002. Aviation Digital Data Service. Adds.awckc.noaa.gov/projects/adds/index.html.

Olson, W., Sarter, N., 1998. As long as i'm in control...: pilot preferences for and experiences with different approaches to automation management. In: Proceedings of the Fourth Annual Symposium on Human Interaction with Complex Systems. pp. 63–72.

Painter, J., Crump, J., Ward, D., 1997. Decision support for the general aviation pilot. In: IEEE International Conference on Systems, Man, and Cybernetics. Vol. 1. pp. 88–93.

Pruyn, P., Greenberg, D., May 1993. Exploring 3D computer graphics in cockpit avionics. IEEE Computer Graphics and Applications , 28–35.

Rich, E., 1998. User modeling via stereotypes. In: Maybury, M., Wahlster, W. (Eds.), Readings in Intelligent User Interfaces. Morgan Kaufmann Publishers, Inc., pp. 329–342.

Ruley, J. D., Summer 2002. Cockpit weather wonders. Pilot Journal , 80–84.

Rumelhart, D., McClelland, J., the PDP Research Group, 1986. Parallel Distributed Processing; Volume 1: Foundations. MIT.

Ruokangas, C. C., Mengshoel, O. J., January 2003. Information filtering using bayesian networks: Effective user interfaces for aviation weather data. In: Proceedings of Intelligent User Interfaces Conference. pp. 280–283.

Russell, S. J., Norvig, P., 1995. Artificial Intelligence: A Modern Approach. Prentice Hall, Inc.

Scanlon, C. H., April 1994. Cockpit graphical weather information shown to enhance efficiency, safety, and situationl awareness. In: Proceedings of the 39th Annual Corporate Aviation Safety Seminar (CASS), Flight Safety Foundation. pp. 83–94.

Spirkovska, L., Lodha, S. K., February 2002. AWE: Aviation weather data visualization environment. Computers and Graphics 26 (1), 169–191.

Spirkovska, L., Lodha, S. K., January 2003. Audio-visual situational awareness for general aviation pilots. In: Proceedings of the SPIE Conference on Visualization and Data Analysis, Santa Clara, CA.

The-Weather-Channel, 2002. Www.weather.com.

Uckun, S., Ruokangas, C., Donohue, P., Tuvi, S., March 1999. AWARE: Technologies for interpreting and presenting aviation weather information. In: Proceedings of the IEEE Aerospace Conference, Snowmass, CO.